

An in silico comparison between margin-based and probabilistic target-planning approaches in head and neck cancer patients

Citation for published version (APA):

Fontanarosa, D., van der Laan, H. P., Witte, M., Shakirin, G., Roelofs, E., Langendijk, J. A., Lambin, P., & van Herk, M. (2013). An in silico comparison between margin-based and probabilistic target-planning approaches in head and neck cancer patients. *Radiotherapy and Oncology*, 109(3), 430-436. <https://doi.org/10.1016/j.radonc.2013.07.012>

Document status and date:

Published: 01/12/2013

DOI:

[10.1016/j.radonc.2013.07.012](https://doi.org/10.1016/j.radonc.2013.07.012)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.



Probabilistic planning

An in silico comparison between margin-based and probabilistic target-planning approaches in head and neck cancer patients



Davide Fontanarosa^{a,b,*}, Hans Paul van der Laan^c, Marnix Witte^d, Georgy Shakirin^e, Erik Roelofs^b, Johannes A. Langendijk^c, Philippe Lambin^{b,1}, Marcel van Herk^{d,1}

^a Philips Research, X-ray Imaging Systems, Eindhoven, The Netherlands; ^b Department of Radiation Oncology (MAASTRO), GROW – School for Oncology and Developmental Biology, Maastricht University Medical Center (MUMC+), Maastricht, The Netherlands; ^c Department of Radiation Oncology, University of Groningen, Groningen, The Netherlands; ^d Department of Radiation Oncology, Netherlands Cancer Institute and Antoni van Leeuwenhoek Hospital (NKI-AVL), Amsterdam, The Netherlands; ^e Philips Technologies, Aachen, Germany

ARTICLE INFO

Article history:

Received 31 August 2012

Received in revised form 12 July 2013

Accepted 24 July 2013

Available online 14 September 2013

Keywords:

Probabilistic planning

Head and neck cancer

Radiotherapy

Margins

ABSTRACT

Background and purpose: To apply target probabilistic planning (TPP) approach to intensity modulated radiotherapy (IMRT) plans for head and neck cancer (HNC) patients.

Material and methods: Twenty plans of HNC patients were re-planned replacing the simultaneous integrated boost IMRT optimization objectives for minimum dose on the boost target and the elective volumes with research probabilistic objectives: the latter allow for explicit handling of systematic and random geometric uncertainties, enabling confidence level based probabilistic treatment planning. Monte-Carlo evaluations of geometrical errors were performed, with endpoints D98%, D2% and Dmean, calculated at a confidence level of 90%. The dose distribution was expanded outside the patient to prevent large bilateral elective treatment volumes ending up in air for probabilistic shifts.

Results: TPP resulted in more regular isodoses and in reduced dose, on average, to organs at risk (OAR), up to more than 6 Gy, while maintaining target coverage and keeping the maximum dose to limiting structures within requirements. In particular, when the surrounding OARs overlap with the planning target volume (PTV) but not with the clinical target volume (CTV), better results were achieved.

Conclusion: The TPP approach was evaluated in HNC patients, and proven to be an efficient tool for managing uncertainties.

© 2013 Elsevier Ireland Ltd. All rights reserved. Radiotherapy and Oncology 109 (2013) 430–436

Head and neck cancer (HNC) patients generally have very complex target volumes, often large and overlapping with or in close proximity to radio-sensitive critical structures [1]. The ability to accurately shape dose in these patients becomes critical [2–4]. With intensity modulated radiotherapy (IMRT), treatments have significantly improved in terms of organs at risk (OARs) sparing while properly covering the target volumes [5]. This could allow for dose escalation strategies on the gross tumor volume (GTV) [6], e.g. using various tracers (such as ¹⁸F]fluoro-deoxy-glucose [7]) to identify the most radioresistant areas inside the GTV [8] and drive the prescription dose distribution accordingly using either a dose painting by contours (DPBC) [9] or a dose painting by numbers (DPBN) [10] approach. For the latter, though, it is inherently impossible to apply standard margin expansions for the regions of interest (ROI) to account for uncertainties. These

can be separated in systematic (Σ) and random (σ) components [11]. The first correspond to the difference between the planning geometry and the average geometry over all treatment fractions. In our study the systematic errors included a systematic component of setup uncertainty and a baseline shift. The latter are the differences between the average treatment geometry and the day-to-day geometries; the random errors refer to errors arising from the positioning of the patient for each fraction. In recent years, many techniques have been proposed to take such uncertainties into account during treatment planning optimization, e.g. probabilistic treatment planning, PP [12–16]. Only our approach, though, allows a confidence level (or probability) based plan optimization, producing results directly comparable to the conventional margin-based approach (for a detailed description of the method see [17]). The purpose for this in silico planning study was to apply target PP (TPP) to HNC cases and to assess if the results obtained were comparable with the traditional margin-based strategies. In case they were not, why and to what extent they differ. This way, also its application to DPBN planning, where no comparison with margin based planning is possible, can be considered reliable.

* Corresponding author. Address: Dr. Tanslaan 12, 6229ET Maastricht, The Netherlands.

E-mail addresses: davide.fontanarosa@maastro.nl, davide.fontanarosa@philips.com (D. Fontanarosa).

¹ The authors contributed equally to this work.

Materials and methods

Patient data

In this study, 20 patients were retrieved from the ROCOCO database (www.mistir.info, [18]) who had undergone primary radiotherapy for HNC and elective or therapeutic treatment of both sides of the neck. The average age was 61 years (range 50–80 years); 20% were female and 80% were male. The primary tumor was located in the oropharynx and hypopharynx in 80% and 20% of the cases, respectively. T-stage was T1 in 13% of the patients, T2 in 20%, T3 in 27% and T4 in 40%; N stage was N0 in 13% of the patients, N1 in 20%, N2a in 7%, N2b in 20% and N2c in 40%.

Radiotherapy treatment planning was performed on a research version of the Philips Pinnacle³ treatment planning system (version 9.100, Philips Radiation Oncology Systems, Fitchburg, WI, USA). Dose was always calculated using the Adaptive Convolve algorithm on computed tomography (CT) scans acquired in treatment position with slice thickness of 2 mm.

Geometrical uncertainties

The original plans, according to the ROCOCO protocol, were created with an expansion CTV to PTV of 0.5 cm in all directions. The hospital values for the standard deviations (SD) of Σ and σ were:

$$\begin{cases} \Sigma_X = 1.11 \text{ mm} \\ \Sigma_Y = 1.10 \text{ mm} \\ \Sigma_Z = 1.04 \text{ mm} \\ \sigma_X = 1.79 \text{ mm} \\ \sigma_Y = 1.52 \text{ mm} \\ \sigma_Z = 1.68 \text{ mm} \end{cases}$$

Patient position deviations were determined from Electronic Portal Images (EPIs) of AP and lateral verification beams with respect to the Digitally Reconstructed Radiographs of these beams. The EPIs were acquired with Elekta iViewGT flat panels before start of treatment. An off-line Shrinking Action Level (SAL) protocol [19] was used for position correction with an initial action level of 6.2 mm (3D vector length) which decreased to 3.1 mm after four fractions. The position deviations of the vertebra were determined. 197 patients were analyzed. The hospital values for Σ and σ were calculated from the patient position deviations with off-line corrections applied according to the method described in the appendix of [20].

These values, according to [21], would produce a margin expansion from CTV to PTV of:

$$\begin{cases} M_X = 4.03 \text{ mm} \\ M_Y = 3.81 \text{ mm} \\ M_Z = 3.77 \text{ mm} \end{cases}$$

This is a margin not comparable to that used for the original plans, which can be explained by the conservative approach of the ROCOCO protocol. But since the errors are then used in the TPP, for a fair comparison they were re-scaled to produce a final margin as close as possible to 0.5 cm.

So the values used were:

$$\begin{cases} \Sigma_X = 1. \text{ mm} \\ \Sigma_Y = 1.42 \text{ mm} \\ \Sigma_Z = 1.34 \text{ mm} \\ \sigma_X = 2.31 \text{ mm} \\ \sigma_Y = 1.96 \text{ mm} \\ \sigma_Z = 2.17 \text{ mm} \end{cases}$$

Which would produce an anisotropic margin with values:

$$\begin{cases} M_X = 5.19 \text{ mm} \\ M_Y = 4.92 \text{ mm} \\ M_Z = 4.87 \text{ mm} \end{cases}$$

Margin based plans

The original plans retrieved were IMRT plans created using the Direct Machine Parameter Optimization (DMPO) module [22], with seven coplanar 6-MV photon beams, at angles: 0°, 50°, 100°, 150°, 210°, 260°, 310°, and a fixed collimator rotation of 5°. To achieve results as general as possible, independent from the specific LINAC used, they were all reoptimized as pure fluence modulated plans, without segmentation.

Organs at risk (OARs), including the parotid glands, submandibular glands, spinal cord, brain stem, optic nerves, optical chiasm, superior and middle pharyngeal constrictor muscle (PCM), and supraglottic larynx, were outlined according to the previously described guidelines [23,24]. The targets in the original plans, the planning target volumes (PTVs), were created according to [25,26]: a PTV54, prescribed with 54.25 Gy in 35 fractions of 1.55 Gy; and a PTV70 prescribed with 70 Gy in 35 fractions of 2.00 Gy. A simultaneous integrated boost (SIB) technique was used where PTV54 was the prophylactic region and PTV70 was the therapeutic region.

Probabilistic planning

In the current work, the concept of PTV was discarded as in TPP the uncertainties are handled without margin expansion. So we created a copy of each of the originally delivered plans. Then we discarded any objective on PTVs or PTV related structures such as ring-like structures around or inside them (MinDose, MaxDose and Uniform Dose objectives in Pinnacle²). We used instead the corresponding CTV54 and CTV70 (mean volumes 127 cm³ (range: 53–327 cm³) and 376 cm³ (range: 254–327 cm³), respectively) as targets. As planning criteria, the ones in the original plans (corresponding to the ROCOCO planning protocol) were used, with the same weights. The same probabilistic coverage of the CTVs (evaluated using the endpoints described in Section ‘Plan evaluation’) was requested on both the original plans and on TPP. The maximum plan dose was 77 Gy and no hotspots (dose exceeding 107% of the prescribed dose) were allowed. After target coverage, the priority was set to not exceed the maximum dose to critical structures (spinal cord, 54 Gy; brainstem, 60 Gy; optic nerves, 54 Gy; and optic chiasm, 54 Gy). Finally, dose to other OARs was minimized as much as possible. In case of overlap between CTV and OAR, the overlapping region was considered as part of the tumor for the optimization process. For each of the newly created plans, the DMPO optimization was switched to Intensity Modulation, no conversion was applied and, after resetting the beams and running the first optimization procedure, two more warm runs (without beams reset) were performed to refine the results.

The TPP plugin provided a research version of all the standard optimization objectives in Pinnacle. We used a research objective named MinDosePP: this is equivalent to the original MinDose objective in Pinnacle, which is met when the region of interest has a minimum dose that is greater than or equal to the target dose. MinDosePP integrates also systematic and random geometric uncertainties during each cost computation. Random errors were simulated by blurring the dose while systematic errors by

² The three objectives are aimed at minimizing the quadratic distance between the minimum, maximum and mean calculated dose and the prescribed dose, respectively.

displacing the ROI with respect to the dose distribution. For every ROI, a cost was calculated for each systematic error, and the probability associated with that error was used to weigh the contribution to the total cost. These contributions were then sorted by cost value. Finally the effect of systematic errors with a certain confidence limit P was calculated adding the cost of every systematic error weighted with its probability until the sum of the probabilities equals P (a more comprehensive description of the method used can be found in [17]). We did not use multiple research objectives on the targets because this would raise correlation problems. In fact, shifts are applied at each cost calculation step and, if multiple objectives independently pick random values from the errors distributions, this would lead to the paradoxical situation of one ROI potentially positioned in two or more different locations with respect to dose distribution at each optimization step. Instead, we used a feature of the plugin for dose painting by contours (named “ROIdelineation”) which allows to set a double prescription value on CTV54 and CTV70 in the same objective, thus avoiding all correlation issues. No research objectives were set on OARs also to avoid correlation issues.

Plan evaluation

The probabilistic evaluation (PE) was performed at the end of each optimization cycle described in Section ‘Probabilistic planning’: if the endpoints described in the following lines matched the values obtained for the CTVs in the original plans, the procedure was stopped. Otherwise it was repeated changing the objective weights until the endpoints matched.

The endpoints were calculated using a standalone software tool (UncertLite, [11]). The tool has the following workflow:

- 1) 10.000 locations of a selected ROI are simulated relative to the dose grid based on given Σ .
- 2) 35 random errors (according to the number of fractions) are sampled from the random error distribution and consecutively added to the systematic error. A cumulative dose is constructed for each systematic error.
- 3) The delivered dose to the selected ROI is calculated for locations defined in step (1) and weighted according to the probability of each location.
- 4) The endpoints D99%, D98%, D2% and Dmean are calculated based on the distribution of CTV doses at a confidence level (CL) of 90%. This CL means that the calculated endpoint values are reached for at least 90% of patients from the population with the given systematic and random errors. Normalization to 35 fractions \times 2 Gy and alpha/beta ratio of 10 Gy was performed.

The CTV54 volumes were smoothed inside Pinnacle using the “smooth contour” tool to produce a comparable volume estimate in both Pinnacle and UncertLite which use slightly different

volume reconstruction methods: CTV54 is typically a very large and complex ROI so it is likely that different results may be produced. Smoothing the contours reduces this effect and the final volumes estimated by Pinnacle and UncertLite were identical.

Both the original and the TPP plans were evaluated with UncertLite for coverage of CTV54 and CTV70.

The dose to OARs was evaluated based on the static planned dose distribution and the reported values were provided by the dose evaluation tool of Pinnacle.

Dose flash in air for evaluation

During plan evaluation, the CTVs were displaced relative to the dose distribution according to systematic and random errors [27] but if they were very close to the skin of the patient they might end up in air. If their shape was wide and bilateral, as for CTV54 in these HNC patients (Fig. 1a), almost every shift on one side or the other would bring the target outside the patient. This resulted in unrealistic dose volume histogram (DVH) calculations for the target. To overcome this issue, during evaluation the dose distribution was expanded around the patient (Fig. 1b). The assumption was that shifting the whole patient and recalculating the dose in the target would give a result comparable to shifting only the target inside the expanded dose distribution and recalculating the DVH (as in [27]). The dose expansion to encompass all possible shifts of the targets was performed by first eroding the dose by 0.7 cm inwards from the external contour of the patient to eliminate the build-up region. This was then followed by a 1.5 cm layered expansion, filling each voxel of the layers from the patient to the expanded contour with the maximum value of dose of the neighboring voxels on the previous layer. A slightly modified version of this approach was already presented in [28].

Results

PTV54 and PTV70 received at least 95% of the prescribed dose on 98% of their volumes for all the original plans and all patients, according to the DVH reported by Pinnacle. Evaluation performed with UncertLite of the original plans showed instead that for CTV54 the coverage was systematically slightly lower than expected, with an average value of 94.1% of prescribed dose to 98% of the volumes. For CTV70, on the other side, the coverage was systematically higher (except for one patient, number 1, where the coverage resulted in only 93.8%), with an average value of 97.7% of prescribed dose to 98% of the volumes. The corresponding TPP plans were matching the original coverage values to an excellent degree (Table 1).

The same coverage on CTVs, when evaluated with UncertLite, was obtained only manually fine tuning the weights of the research objectives on TPP plans. No threshold was set a priori on this variation, as they were only intended to reproduce as

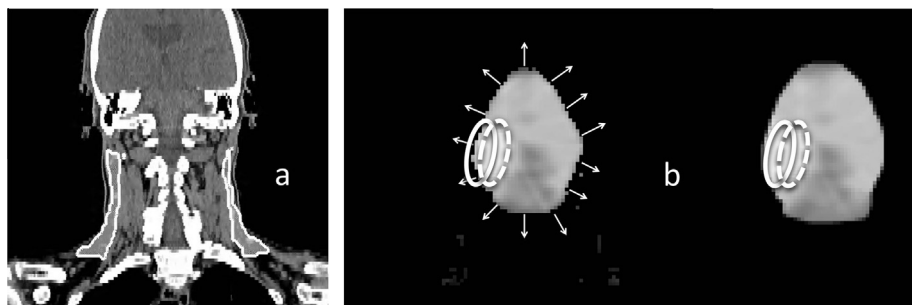


Fig. 1. (a) CTV54 in a coronal projection showing its proximity to skin on both sides of the neck. (b) The original dose distribution as retrieved from the original Pinnacle plan (on the left) is expanded around the patient (on the right) enough to encompass all possible probabilistic shifts of the targets (dashed before the shift and solid after).

Table 1

Dose to 98% of the volumes of CTV54 (on the left) and CTV70 (on the right) as absolute values and as percent of the prescribed dose (54.25 Gy to CTV54 and 70 Gy to CTV70), according to the evaluation performed with UncertLite. Average values and standard deviations are also reported, showing an excellent agreement between CTVs coverage in the original plans and the PP ones.

Patient	CTV54				CTV70			
	Original		PP		Original		PP	
	Dose (Gy)	V98 (%)	Dose (Gy)	V98 (%)	Dose (Gy)	V98 (%)	Dose (Gy)	V98 (%)
1	50.61	93.3	50.84	93.7	65.64	93.8	65.29	93.3
2	51.09	94.2	50.91	93.8	68.78	98.3	68.90	98.4
3	51.40	94.7	51.19	94.4	68.68	98.1	68.57	98.0
4	51.18	94.3	51.78	95.4	68.41	97.7	67.83	96.9
5	51.04	94.1	51.21	94.4	68.86	98.4	68.85	98.4
6	51.36	94.7	51.55	95.0	68.84	98.3	69.03	98.6
7	50.67	93.4	51.44	94.8	67.36	96.2	67.96	97.1
8	51.30	94.6	51.45	94.8	69.08	98.7	69.08	98.7
9	51.38	94.7	51.44	94.8	68.66	98.1	68.71	98.2
10	51.12	94.2	51.22	94.4	68.35	97.6	67.55	96.5
11	51.03	94.1	51.24	94.5	68.48	97.8	68.38	97.7
12	51.15	94.3	51.22	94.4	68.67	98.1	68.33	97.6
13	50.93	93.9	50.97	94.0	68.61	98.0	68.73	98.2
14	51.33	94.6	50.97	94.0	68.57	98.0	67.84	96.9
15	51.27	94.5	51.57	95.1	68.55	97.9	68.75	98.2
16	51.47	94.9	51.24	94.5	68.79	98.3	68.73	98.2
17	51.16	94.3	51.33	94.6	69.16	98.8	68.75	98.2
18	51.22	94.4	51.71	95.3	68.43	97.8	68.62	98.0
19	50.27	92.7	50.91	93.8	68.32	97.6	68.33	97.6
20	50.42	92.9	50.89	93.8	67.25	96.1	67.93	97.0
Average	51.07	94.1	51.25	94.5	68.37	97.7	68.31	97.6
SD	0.33	0.6	0.28	0.5	0.80	1.1	0.83	1.2

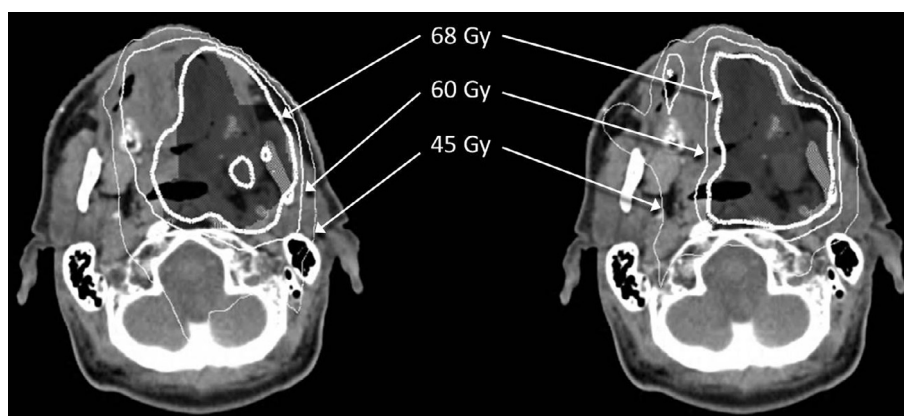


Fig. 2. On the left, one of the IMRT plans with PTV70 colorwashed in dark gray; the lines correspond to the different isodose levels 45, 60 and 68 Gy. On the right, the same case planned with the probabilistic planning strategy. A better conformity and steeper gradients can be seen, resulting in improved OARs sparing.

accurately as possible the original CTV coverage. The optimization time, with the hardware available (Philips 810x workstation), was on average 9 min.

In most patients, TPP produced more regular and smoother isodoses. This is probably the result of the incorporation of systematic uncertainties which, during optimization, produce a shift of dose volumes and ROIs relative to each other: for each shift the discrepancy between prescription and calculated dose is minimized. Thus the dose peaks are smoothed more efficiently and hot and cold spots are reduced (Fig. 2).

Replacing the objectives on PTVs in the original plans with a combined MinDosePP objective on CTV54 and CTV70 (as described in Section 'Probabilistic planning'), and using the original weights (or allowing for minimum variations around it on the research objective to reproduce CTV coverage as described in the previous paragraph) for the associated cost function, the maximum dose to limiting structures was kept within requirements also in the

TPP plans. The dose to OARs, instead, was reduced on average in all patients for all the considered ROIs (Table 2).

The maximum dose reduction observed with TPP was up to more than 6 Gy (for one of the submandibular glands, Table 2). The mean dose to healthy tissues (patient minus CTVs) also remained constant.

Discussion

We found that, with TPP, increased sparing of OARs could be obtained relative to margin based planning, while the evaluation procedure showed a similar probability of effective CTV coverage for both methods. In this study, the TPP approach showed beneficial when it was applied to both CTV54 and CTV70, however it might be specifically required in the application of DPBN. In that case, TPP could be very useful to boost sub volumes of CTV70 that have high positron emission tomography (PET) standard uptake values

Table 2

Average values for the twenty HNC patients of dose and dose variation (difference between the dose to the ROI in PP and the dose to the ROI in the original plan) to limiting structures and OARs between the original plan and the PP over the whole treatment course. Min and max values are also reported, which show that only minor variations unfavorable to PP (negative values) have occurred while the advantages (positive values) were important for some structures of some patients. In the last row, mean dose to healthy tissues is reported. The last four columns show the percentage of patients obtaining a max dose variation between the original plan and the PP larger than 1, 2, 3 and 5 Gy respectively for each reported ROI. The dose to OARs was evaluated based on the static planned dose distribution.

OAR type	OAR	Population mean dose in the original plan (Gy)	Population mean dose in PP (Gy)	Population mean dose variation (Gy)	Min dose in population for the original plan (Gy)	Min dose in population for PP (Gy)	Min dose in population variation (Gy)	Max dose in population for the original plan (Gy)	Max dose in population for PP (Gy)	Max dose in population variation (Gy)	Percentage of patients with max dose variation larger than			
											1 Gy	2 Gy	3 Gy	5 Gy
Limiting structures (max dose)	Brainstem	49.56	48.62	0.94	35.01	35.61	−0.60	54.27	51.09	3.18	55	20	5	0
	Spinal cord	43.94	43.39	0.55	46.26	46.62	−0.36	41.28	38.77	2.51	15	10	0	0
Salivary glands (mean dose)	Parotid ipsilateral	44.00	42.70	1.30	48.25	48.24	0.01	44.68	38.79	5.89	55	20	15	5
	Parotid contralateral	33.46	32.89	0.57	25.06	25.12	−0.06	48.75	46.54	2.21	20	5	0	0
	Submand. ipsilateral	67.36	66.46	0.90	69.94	69.96	−0.02	68.04	61.99	6.05	25	5	5	5
	Submand. contralateral	59.89	59.01	0.88	47.69	48.17	−0.48	69.42	67.09	2.33	35	5	0	0
	Swallowing OARs (mean dose)	64.65	63.75	0.90	69.60	70.20	−0.60	60.11	58.35	1.76	45	0	0	0
Swallowing OARs (mean dose)	PCM	62.40	61.28	1.12	49.53	49.73	−0.20	68.50	64.12	4.38	45	20	5	0
	Middle PCM	59.72	58.68	1.04	69.33	69.83	−0.50	70.86	66.92	3.94	40	20	5	0
	Supraglottic larynx	59.72	58.68	1.04	69.33	69.83	−0.50	70.86	66.92	3.94	40	20	5	0
Healthy tissues (mean dose)		12.74	12.42	0.32	10.10	10.07	0.03	15.47	14.80	0.67	0	0	0	0

(SUV) [6,8]. When the use of TPP is restricted to CTV70, due to its limited extent and above all to the relatively small dose difference between the prescription on CTV54 and CTV70, it is very difficult to achieve substantial sparing on OARs surrounding CTV70 (and typically at least partially included in CTV54) and results produced by TPP are likely to match those with margin-based planning.

Our results suggest that on average it is unlikely to achieve important dose reductions to OARs when TPP is applied to targets in HNC cases, but it is possible that specific patients benefit considerably more than others. Typically, a good result is achieved on an OAR when its contour overlaps with the original PTV of the target but not with the CTV. In fact, in these cases, TPP producing steeper isodose lines around the CTV can spare dose to the OAR. Typically CTVs are surrounded by many OARs and the previous statement cannot hold true for all of them simultaneously. However, for planning strategies aiming at sparing specific OARs (like swallowing structures, for example), the overlap can be assessed for only those structures and, if the requirement is satisfied, TPP can be considered to reduce dose to those organs. Moreover, it must be noted that these results on OARs are also connected to our strict focus on CTV coverage: as explained in Section ‘Probabilistic planning’, the overlapping areas between targets and OARs were always considered by the optimizer as targets. So this work is presenting a worst case scenario, whereas adopting a less aggressive approach might result in a generally better sparing of healthy tissues.

As explained in Section ‘Probabilistic planning’, TPP was not applied on OARs because it was not possible to correlate the shifts among the different research objectives. In inverse planning, it is extremely complicated to take the correlation into account because the objectives are independently set. Nonetheless, it is advisable in the future to extend the PP to OARs, because small rotations result in small differences at the isocenter, where the uncertainties are typically calculated, whereas become important the larger is the distance from it: so OARs located far from the target can possibly show overdosage due to very small rotations. Dose is expected to decrease rapidly moving away from the target, so this effect is not likely to be important, in general. But for HNC patients where

targets can extend considerably far from the isocenter, and thus also steep dose gradients, this effect might be important. It would be interesting also to apply PE on OARs, and it is part of our future work: a systematic assessment of the impact of the uncertainties on dose to OARs as a function of their magnitude and of the specific OAR. Since there is now general consensus on the importance of local variability of uncertainties for HNC patients [29–31], it would be of extreme interest to assess how each specific structure is sensitive to this variation. In this specific work, the application of PE on OARs was not feasible because all the dose limits set on the original plans and by the ROCOCO protocol were set on OARs without expansion. Therefore, the comparison between the ROIs with margin expansion in the original plans with the ROIs with probabilistic shifts in the PP plans was not possible as we did for targets (where the original dose constraints were set on the PTV while in TPP they were set on the CTVs).

PP and PE applied to large targets (as is the case for the CTV54 presented here) in principle suffer from the same issues, because the errors reported by the hospitals are calculated typically at the beam isocenter: how to extend this information to surrounding regions is in general unknown and patient-dependent for HNC [31]. For example, head flexion with respect to the base of the neck, tumor or parotid shrinkage or swallowing can create relatively small differential displacements of the tissues which can in turn result in dramatically different dose delivery: here in fact the large targets require high dose levels in large areas. Therefore, in general, errors have to be known very accurately, in particular in terms of local variations, otherwise important dose misestimates might occur. It would be interesting to assess whether a population-based uncertainty calculation is sufficient for these types of patients or it is necessary to evaluate every single case separately. Paradoxically, small targets in areas undergoing much larger movements than the head and neck region, such as non-small cells lung cancers (NSCLC), require dose distributions less extended thus making accuracy in tracking OARs positions less important. Moreover, in lung cancer patients, respiratory motion can nowadays be predicted using 4DCT datasets [26], and deformation algorithms could

provide reliable estimates of error propagation patterns from the target to every other tissue inside the patient.

In this work we assumed that the dose cloud approach is valid, and that shifting the ROIs inside a static dose distribution is reasonably similar to recalculating the dose after the shift. This assumption was based on [28] where a similar approach as our dose flash in air (there named Corrected Convolution) was implemented, showing acceptable error levels also close to the surface. The advantage of our approach, where the shifts are produced in the dose space instead of fluence space, is in the calculation times: in the first case, only the DVH must be computed, in the second for each shift also a full dose re-calculation is required.

The approach we used was to reproduce on the TPP plans with the highest possible accuracy the values of the endpoints for CTVs of the original plans. Then we evaluated the doses on the OARs using the standard tools provided by Pinnacle since, as discussed previously, no extension of PE to structures other than targets was feasible. So this can be considered a mixed approach, since a proper evaluation should be performed with Monte Carlo methods on all ROIs.

This work is based on a group of patients part of another study previously published [32], where sufficient coverage of both PTV54 and PTV70 was reported. Our work shows that Monte-Carlo evaluation can identify cases where the real coverage is not as expected (see Table 1). It is necessary to highlight that CTV to PTV expansion in the original plans was restricted to 0.5 cm from the skin to exclude buildup region: so in this shallow area a prescription to TPP CTVs and to the original PTVs do not aim at covering the same area.

PP was previously applied to one HNC case [27], but with DPBN approach, where no comparison with standard margins was possible and no extensive study of the characteristics of this site was yet performed. Also, different methods were tested on HNC cases (see for example [14]), but only the method used in this work implements a number of displacements sufficient to statistically account for positioning errors. This allows an optimization with a given confidence of a physical dose parameter, reproducing what is requested with standard margin-based planning: thus a direct comparison with conventional planning strategies is possible, as opposite to the other approaches (like the previously cited MIGA [14], or based on coverage probabilities [12], or performing dose error simulations prior to cost values computation [15]). Our implementation of the optimization objectives is similar to the coverage optimized approach introduced in [13], but adapted to the Pinnacle environment.

The peculiarities of TPP applied to HNC cases were specifically addressed in this work for the first time: very high and complicated dose gradients, multiple extended complex targets and their proximity to patient's surface require solutions, presented here, different than what is described elsewhere. For example, where prostate cases are treated, as in [17] (where also rotations are considered) or [33] (where fractionation effect was not considered). In these works, toxicity reduction thanks to PP was reported.

TPP provides on average a better sparing of OAR tissues, reducing the overlap area between the irradiated volume and the surrounding structures. Categorizing plans according to OAR overlap with PTV but not CTV potentially allows for an a priori assessment of TPP strategy advantage over traditional margin based approach. The dose to OARs was decreased without putting further efforts on optimization: everything was kept as in the original plans, only one objective was replaced by the research version, and minimal weight changes were necessary for it to reproduce similar endpoints (as described in Section 'Probabilistic planning'). This means the decrease was due only to novel planning strategy. The benefit of TPP is very dependent on individual patient factors, as is illustrated by the ranges in Table 1. More cases are being investigated

within the ROCOCO framework, allowing for comparison with other treatment modalities.

Conflicts of interest

No actual or potential conflicts of interest exist.

Acknowledgments

The authors wish to thank Dr. Matthieu Bal from Philips Healthcare, Eindhoven, The Netherlands for providing the research version of Pinnacle and supporting the plugin development. This research was supported by the Center for Translational Molecular Medicine AIRFORCE project.

References

- [1] Gregoire V, Coche E, Cosnard G, Hamoir M, Reyckler H. Selection and delineation of lymph node target volumes in head and neck conformal radiotherapy. Proposal for standardizing terminology and procedure based on the surgical experience. *Radiother Oncol* 2000;56:135–50.
- [2] Claus F, De Gersem W, De Wagter C, et al. An implementation strategy for IMRT of ethmoid sinus cancer with bilateral sparing of the optic pathways. *Int J Radiat Oncol Biol Phys* 2001;51:318–31.
- [3] Eisbruch A, Ten Haken RK, Kim HM, Marsh LH, Ship JA. Dose, volume, and function relationships in parotid salivary glands following conformal and intensity-modulated irradiation of head and neck cancer. *Int J Radiat Oncol Biol Phys* 1999;45:577–87.
- [4] Lee N, Xia P, Quivey JM, et al. Intensity-modulated radiotherapy in the treatment of nasopharyngeal carcinoma: an update of the UCSF experience. *Int J Radiat Oncol Biol Phys* 2002;53:12–22.
- [5] Eisbruch A, Foote RL, O'Sullivan B, Beitler JJ, Vikram B. Intensity-modulated radiation therapy for head and neck cancer: emphasis on the selection and delineation of the targets. *Semin Radiat Oncol* 2002;12:238–49.
- [6] Dawson LA, Anzai Y, Marsh L, et al. Patterns of local-regional recurrence following parotid-sparing conformal and segmental intensity-modulated radiotherapy for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2000;46:1117–26.
- [7] Daisne JF, Duprez T, Weynand B, et al. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. *Radiology* 2004;233:93–100.
- [8] Chao KS, Ozyigit G, Tran BN, Cengiz M, Dempsey JF, Low DA. Patterns of failure in patients receiving definitive and postoperative IMRT for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2003;55:312–21.
- [9] Meijer G, Steenhuijsen J, Bal M, De Jaeger K, Schuring D, Theuvs J. Dose painting by contours versus dose painting by numbers for stage II/III lung cancer: practical implications of using a broad or sharp brush. *Radiother Oncol* 2011;100:396–401.
- [10] Vanderstraeten B, Duthoy W, De Gersem W, De Neve W, Thierens H. [18F]fluoro-deoxy-glucose positron emission tomography ([18F]FDG-PET) voxel intensity-based intensity-modulated radiation therapy (IMRT) for head and neck cancer. *Radiother Oncol* 2006;79:249–58.
- [11] van Herk M, Remeijer P, Lebesque JV. Inclusion of geometric uncertainties in treatment plan evaluation. *Int J Radiat Oncol Biol Phys* 2002;52:1407–22.
- [12] Baum C, Alber M, Birkner M, Nusslin F. Robust treatment planning for intensity modulated radiotherapy of prostate cancer based on coverage probabilities. *Radiother Oncol* 2006;78:27–35.
- [13] Gordon JJ, Sayah N, Weiss E, Siebers JV. Coverage optimized planning: probabilistic treatment planning based on dose coverage histogram criteria. *Med Phys* 2010;37:550.
- [14] McShan DL, Kessler ML, Vineberg K, Fraass BA. Inverse plan optimization accounting for random geometric uncertainties with a multiple instance geometry approximation (MIGA). *Med Phys* 2006;33:1510–21.
- [15] Unkelbach J, Oelfke U. Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions. *Phys Med Biol* 2004;49:4005–29.
- [16] Witte MG, van der Geer J, Schneider C, Lebesque JV, van Alber M, Herk M. IMRT optimization including random and systematic geometric errors based on the expectation of TCP and NTCP. *Med Phys* 2007;34:3544.
- [17] Bohoslavsky R, Witte MG, van Janssen TM, Herk M. Probabilistic objective functions for margin-less IMRT planning. *Phys Med Biol* 2013;58:3563–80.
- [18] Roelofs E, Persoon L, Qamhiyeh S, et al. Design of and technical challenges involved in a framework for multicentric radiotherapy treatment planning studies. *Radiother Oncol* 2010;97:567–71.
- [19] Bel A, van Herk M, Bartelink H, Lebesque JV. A verification procedure to improve patient set-up accuracy using portal images. *Radiother Oncol* 1993;29:253–60.
- [20] de Boer HC, van Sornsens de Koste JR, Senan S, Visser AG, Heijmen BJ. Analysis and reduction of 3D systematic and random setup errors during the simulation and treatment of lung cancer patients with CT-based external beam radiotherapy dose planning. *Int J Radiat Oncol Biol Phys* 2001;49:857–68.

- [21] van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose–population histograms for deriving treatment margins in radiotherapy. *Int J Radiat Oncol Biol Phys* 2000;47:1121–35.
- [22] Siebers JV, Lauterbach M, Keall PJ, Mohan R. Incorporating multi-leaf collimator leaf sequencing into iterative IMRT optimization. *Med Phys* 2002;29:952–9.
- [23] Christianen ME, Langendijk JA, Westerlaan HE, van de Water TA, Bijl HP. Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. *Radiother Oncol* 2011;101:394–402.
- [24] van de Water TA, Bijl HP, Westerlaan HE, Langendijk JA. Delineation guidelines for organs at risk involved in radiation-induced salivary dysfunction and xerostomia. *Radiother Oncol* 2009;93:545–52.
- [25] Gregoire V, Eisbruch A, Hamoir M, Levendag P. Proposal for the delineation of the nodal CTV in the node-positive and the post-operative neck. *Radiother Oncol* 2006;79:15–20.
- [26] Gregoire V, Levendag P, Ang KK, et al. CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC, RTOG consensus guidelines. *Radiother Oncol* 2003;69:227–36.
- [27] Witte M, Shakirin G, Houweling A, van Peulen H, Herk M. Dealing with geometric uncertainties in dose painting by numbers: introducing the ΔVH . *Radiother Oncol* 2011;100:402–6.
- [28] Craig T, Battista J, Dyk JV. Limitations of a convolution method for modeling geometric uncertainties in radiation therapy. I. The effect of shift invariance. *Med Phys* 2003;30:2001.
- [29] Giske K, Stoiber EM, Schwarz M, et al. Local setup errors in image-guided radiotherapy for head and neck cancer patients immobilized with a custom-made device. *Int J Radiat Oncol Biol Phys* 2011;80:582–9.
- [30] van Kranen S, van Beek S, Rasch C, van Herk M, Sonke JJ. Setup uncertainties of anatomical sub-regions in head-and-neck cancer patients after offline CBCT guidance. *Int J Radiat Oncol Biol Phys* 2009;73:1566–73.
- [31] Birkner M, Thorwarth D, Poser A, Ammazalorso F, Alber M. Analysis of the rigid and deformable component of setup inaccuracies on portal images in head and neck radiotherapy. *Phys Med Biol* 2007;52:5721–33.
- [32] van der Laan HP, Christianen ME, Bijl HP, Schilstra C, Langendijk JA. The potential benefit of swallowing sparing intensity modulated radiotherapy to reduce swallowing dysfunction: an in silico planning comparative study. *Radiother Oncol* 2012;103:76–81.
- [33] Moore JA, Gordon JJ, Anscher M, Silva J, Siebers JV. Comparisons of treatment optimization directly incorporating systematic patient setup uncertainty with a margin-based approach. *Med Phys* 2012;39:1102–11.